

1 **ABSTRACT**

2 Automatic detection and tracking of multiple individuals includes receiving
3 a frame of video and/or audio content and identifying a candidate area for a new
4 face region in the frame. One or more hierarchical verification levels are used to
5 verify whether a human face is in the candidate area, and an indication made that
6 the candidate area includes a face if the one or more hierarchical verification
7 levels verify that a human face is in the candidate area. A plurality of audio and/or
8 video cues are used to track each verified face in the video content from frame to
9 frame.

10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25